UNIVERSIDADE FEDERAL DE SERGIPE
DECAT DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS

Processos estocásticos

Disciplina ofertada pelo DECAT/UFS

Código: ESTAT0077
Nível: Graduação
Carga horária: 60h
Período: 2020.2

Professor responsável e ministrante: Luiz Henrique Dore

Aula 2: definição, classificação e distribuição de um processo estocástico

Sumário

Metas

1. Introduzir o aluno ou aluna ao estudo dos processos estocásticos, apresentando uma definição para tais processos, exemplos.
2. Mostrar como classificar os processos estocásticos.
3. Apresentar noções distribuições de probabilidade no contexto de processos estocásticos.

Objetivos

- Após estudar essa aula, o aluno ou aluna será capaz de:
  1. conceitualizar processos estocásticos e explicar a sua utilidade e importância no desenvolvimento de métodos estatísticos;
  2. apresentar exemplos de processos estocásticos e de fenômenos que podem ser modelados como um processo estocástico;
  3. descrever o espaço de estados e o conjunto de índices de um processo estocástico e classificar o processo de acordo com essas duas características.
  4. apresentar uma noção sobre a distribuição de probabilidade de ordem k de um processo estocástico;
  5. definir as funções de distribuição de ordem k, de probabilidade de ordem k e a de densidade de probabilidade de ordem k;
  6. em situações simples, obter a distribuição de probabilidade de um processo estocástico.

Pré-requisitos

1. Não há pré-requisitos.

Introdução

- A estatística é uma ciência que lida com desenvolvimento de teorias e métodos matemáticos e computacionais com o objetivo de compreender o comportamento de uma entidade, a partir de observações dessa entidade.
- A entidade é representada por características mensuráveis, as quais são denominadas variáveis.
- Cada vez que a entidade é observada, as variáveis são medidas e seus valores são registrados.
- Em geral, os valores referentes a diferentes observações da entidade são diferentes entre si.
- Dessa forma, o conjunto dos valores registrados exibe uma variabilidade.

Introdução

- Essa variabilidade é analisada a partir da frequência com a qual os valores ocorrem. O objetivo é identificar padrões de frequência no conjunto dos valores registrados que permitam responder questões a respeito do comportamento da entidade.
- Métodos matemáticos e computacionais, com fundamentos na teoria da probabilidade, são utilizados para esse fim.
- Tais métodos são denominados métodos estatísticos.
- A premissa básica para o funcionamento de métodos estatísticos é a de que a variabilidade observada no conjunto de valores registrados é regida por algum mecanismo aleatório.
- Os métodos estatísticos incorporam essa premissa adotando a suposição de que **os valores registrados são realizações de variáveis aleatórias**.

Introdução

- Seja Y a variável que representa a entidade.
- Pretende-se estudar a entidade a partir de n observações de Y.
- Sejam $Y_1, \cdots, Y_n$ variáveis que representam os valores de Y **a serem observados**.
- Isto é, para cada $i \in \{1, \cdots, n\}$, Y_i representa um valor que **só será conhecido no momento em que o i-ésimo valor da variável Y for observado e registrado**
- Os métodos estatísticos assumem as variáveis $Y_1, \cdots, Y_n$ são **variáveis aleatórias**.
- Em outras palavras, os métodos estatísticos partem da suposição de que **os valores de Y a serem observados formam uma sequência de variáveis aleatórias**.

Introdução

- Muitos métodos estatísticos utilizados na prática, assumem que as variáveis aleatórias nessa sequência são **independentes**.
- Isto é, assumem que $Y_1, \cdots, Y_n$ é uma sequência de variáveis aleatórias independentes.
- Porém, com frequência, essas variáveis exibem algum tipo de interdependência, o que torna essa suposição irrealística.
- Por exemplo, em geral, os valores de $Y_1, \cdots, Y_n$ são observados em diferentes instantes de tempo.
- É comum tais valores exibirem **dependência temporal**.
- Isso significa que, **o conhecimento dos valores da variável em instantes de tempo passados, altera o conhecimento sobre o comportamento da variável no futuro**.

Introdução

- Os valores de $Y_1, \cdots, Y_n$ podem ser observados em diferentes posições no espaço.
- Tais valores podem vir a exibir uma **dependência espacial**.
- Essa dependência espacial se manifesta da seguinte maneira: **o conhecimento dos valores da variável numa dada região do espaço altera o conhecimento sobre o comportamento da variável em regiões vizinhas**.
- Um processo estocástico é uma **coleção de variáveis aleatórias que podem (ou não) ser interdependentes**.
- No estudo de um processo estocástico, métodos matemáticos e, mais especificamente, probabilísticos, são usados com o intuito de compreender o comportamento das variáveis aleatórias na coleção que define o processo.

Introdução

- O conhecimento sobre processos estocásticos pode ser utilizado no desenvolvimento métodos estatísticos que admitem alguma estrutura de dependência entre $Y_1, \cdots, Y_n$.
- A presente aula tem como objetivo introduzir o aluno ou aluna ao estudo dos processos estocásticos.
- O tópico 2 apresenta uma definição para processos estocásticos e alguns exemplos.
- As variáveis aleatórias num processo estocástico são unicamente identificadas por um **índice**.
- O tópico 3 apresenta uma forma de classificar um processo estocástico, com relação aos valores das variáveis aleatórias que o compõem e com relação aos **índices** que as identificam.

Introdução

- Estudar uma variável aleatória significa estudar sua distribuição de probabilidade.
- Isso também vale para um processo estocásticos.
- No tópico 4, são introduzidas noções relacionadas à distribuição de probabilidade de um processo estocástico, incluindo a noção de distribuição de ordem k de um processo.

Definição de um processo estocástico

Definição 1

Seja T um subconjunto qualquer dos números reais. Um processo estocástico é um conjunto $\{X(t), t \in T\}$ tal que, para cada $t \in T$, $X(t)$ é uma variável aleatória [2, p. 47].

- Um processo estocástico é um conjunto de variáveis aleatórias unicamente identificadas por um **índice** t.
- O conjunto $T \subset \mathbb{R}$ é o **conjunto de índices**.
- Na prática, a variável aleatória $X(t) \in \{X(t), t \in T\}$ costuma representar o resultado de um experimento aleatório, realizado num instante de tempo t.
- Dessa forma, o índice t é comumente referido como **tempo**.

Definição de um processo estocástico

- Nesse caso, o conjunto de índices costuma ser $T = \{0, 1, 2, \cdots\}$ ou o intervalo $T = [0, \infty)$. Se $t \in \{0, 1, \cdots\}$, o processo estocástico é escrito como $\{X_n, n = 0, 1, \cdots\}$.

Exemplo 1

Uma partícula pode se movimentar para esquerda e para a direita. No tempo 0, a partícula se encontra na origem. A cada unidade de tempo, uma moeda é arremessada. Se o resultado é coroa, a partícula se move uma unidade para a direita. Se o resultado é cara, a partícula se move uma unidade para a esquerda. Seja X_n a variável aleatória que representa a posição da partícula após n arremessos. O conjunto $\{X_n, n = 0, 1, \cdots\}$ é um processo estocástico conhecido como **passeio aleatório**.

Definição de um processo estocástico

Exemplo 2

Seja X_n a variável aleatória que representa o preço médio de uma ação, em um dia n. O conjunto $\{X_n, n = 0, 1, \cdots\}$ é um processo estocástico.

Exemplo 3

Seja X_n a variável aleatória que representa o número de itens defeituosos, encontrados num determinado dia n, em uma linha de produção. Nesse caso, $\{X_n, n = 0, 1, 2, \cdots\}$ é um processo estocástico.

Definição de um processo estocástico

Exemplo 4

Suponha que temos duas caixas (1 e 2) e d bolas, numeradas de 1 até d. Inicialmente, algumas dessas bolas se encontram na caixa 1 e as demais na caixa 2. Um inteiro de 1 a d é selecionado aleatoriamente e a bola correspondente é removida da caixa à qual ela pertence e posta na outra caixa. Esse procedimento é repetido, indefinidamente, de modo que as seleções sejam independentes. Seja X_n o número de bolas na caixa 1, após n seleções. O conjunto $\{X_n, n = 0, 1, \cdots\}$ é um processo estocástico conhecido como **cadeia de Ehrenfest**. A cadeia de Ehrenfest é um modelo simples de transferência de calor ou de moléculas de gás entre dois corpos isolados [1, p. 7].

Definição de um processo estocástico

Exemplo 5

Suponha que um jogador (apostador) possui, inicialmente, um certo capital em reais. Ele resolve fazer uma série de apostas de um real contra a casa. Ele ganha a aposta com probabilidade p e perde com probabilidade $1 - p$. Quando o seu capital chega a zero, ele está arruinado (falido) e, uma vez arruinado, ele permanece arruinado. Isto é, uma vez arruinado, seu capital não evolui mais e permanece zero. Seja X_n o capital do jogador na n-ésima jogada. O conjunto $\{X_n, n = 0, 1, \cdots\}$ é um processo estocástico conhecido como **ruína do jogador** [1, p. 8].

Definição de um processo estocástico

Exemplo 6

Em um determinado ponto, um sensor registra graficamente a temperatura por um período indefinido de tempo. Seja $X(t)$ a variável aleatória que representa a temperatura no instante de tempo t. Então, $\{X(t), t \geq 0\}$ é um processo estocástico.

- Nos exemplos anteriores, o índice t representou o tempo.
- Além do tempo, o índice t, no processo $\{X(t), t \geq 0\}$, pode representar a **posição** onde a variável aleatória $X(t)$ é observada.
- Nesse caso, o conjunto de índices T representa as possíveis posições e o processo estocástico representa os resultados de um experimento aleatório realizado nas várias posições possíveis.

Definição de um processo estocástico

Exemplo 7

A temperatura num determinado ponto é medida em diferentes alturas, até uma altura máxima L. Seja $X(h)$ a variável aleatória que representa a temperatura na altura h. Então, o conjunto $\{X(h), 0 \leq h \leq L\}$ é um processo estocástico.

Definição 2

Se T é finito ou infinito enumerável, então $\{X(t), t \in T\}$ é dito ser um processo com tempo discreto; se T é infinito não enumerável, então $\{X(t), t \in T\}$ é dito ser um processo com tempo contínuo [2, p. 47].

Classificação de um processo estocástico

Definição 3

O conjunto $S_{X(t)}$, formado pelos valores que as variáveis aleatórias $X(t)$ podem assumir, é denominado espaço de estados do processo estocástico $\{X(t), t \in T\}$; se $S_{X(t)}$ é finito ou infinito enumerável, então $\{X(t), t \in T\}$ é dito ser um processo com espaço de estados discreto; se $S_{X(t)}$ é infinito não enumerável, então $\{X(t), t \in T\}$ é dito ser um processo com espaço de estados contínuo [2, p. 48].

- Processos estocásticos podem ser classificados, com relação ao conjunto de índices e ao espaço de estados, de quatro maneiras.

Classificação de um processo estocástico

- Processos estocásticos pode ser um processo com 1) espaço de estados discreto e tempo discreto; 2) espaço de estados discreto e tempo contínuo; 3) espaço de estados contínuo e tempo discreto; 4) espaço de estados contínuo e tempo contínuo.

Exemplo 8

Um exemplo elementar de um processo estocástico com tempo contínuo é o processo $\{X(t), t \geq 0\}$ definido como

$$X(t) = Yt, \quad \forall t \geq 0,$$

onde Y é uma variável aleatória tendo uma distribuição qualquer [2, p. 48]. O espaço de estados desse processo depende de Y ser identicamente nula ou não.

Classificação de um processo estocástico

Continuação do exemplo 8

Se $Y \equiv 0$, então $S_{X(t)} = \{0\}$. Nesse caso, o processo tem espaço de estados discreto e tempo contínuo. Se $Y < 0$, então $S_{X(t)} = (-\infty, 0]$. Se $Y > 0$, então $S_{X(t)} = [0, \infty)$. Nesses dois últimos casos, os processos têm espaço de estados contínuo e tempo contínuo.

Exemplo 9

O espaço de estados do passeio aleatório, no exemplo 1, é

$$S_{X_n} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}.$$

Portanto, esse passeio aleatório é um processo com espaço de estados discreto e tempo discreto.

Classificação de um processo estocástico

Exemplo 10

O espaço de estados da cadeia de Ehrenfest, no exemplo 4, é

$$S_{X_n} = \{0, 1, 2, \cdots, d\}.$$

Portanto, a cadeia de Ehrenfest é um processo com espaço de estados discreto e tempo discreto.

Exemplo 11

O espaço de estados do preço da ação, no exemplo 2, é

$$S_{X_n} = [0, \infty),$$

Portanto, esse processo possui espaço de estados contínuo e tempo discreto.

Distribuição de ordem k

Definição 4

[2, p. 49] A função de distribuição de ordem k do processo estocástico $\{X(t), t \in T\}$, nos tempos $t_1, \cdots, t_k$, é a função de distribuição conjunta do vetor aleatório $(X(t_1), \cdots, X(t_k))$:

$$F(x_1, \cdots, x_k; t_1, \cdots, t_k) = P[X(t_1) \leq x_1, \cdots, X(t_k) \leq x_k].$$

Definição 5

[2, p. 49] Se o processo estocástico $\{X(t), t \in T\}$ tem espaço de estados discreto, então define-se a função de probabilidade ordem k desse processo, nos tempos $t_1, \cdots, t_k$, como

$$p(x_1, \cdots, x_k; t_1, \cdots, t_k) = P[X(t_1) = x_1, \cdots, X(t_k) = x_k].$$

Distribuição de ordem k

Definição 6

[2, p. 49] Se o processo estocástico $\{X(t), t \in T\}$ tem espaço de estados contínuo, então define-se a sua função de densidade de probabilidade de ordem k, nos tempos $t_1, \cdots, t_k$, como

$$f(x_1, \cdots, x_k; t_1, \cdots, t_k) = \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1, \cdots, x_k; t_1, \cdots, t_k).$$

- A distribuição de probabilidade de ordem k é a distribuição de probabilidade conjunta do processo em k instantes de tempo.
- Quando $k = 1$, obtém-se a distribuição do processo em um único instante de tempo.

Distribuição de ordem k

Exemplo 12

Seja $\{X_n, n = 0, 1, \cdots\}$ um processo estocástico com espaço de estados $S_{X(t)} = \{0, 1, 2, \cdots\}$. Utilizando a notação introduzida na definição 5, pode-se escrever a probabilidade do processo assumir o valor 5 no tempo 1 da seguinte forma:

$$P[X_1 = 5] = p(5; 1)$$

A probabilidade do processo assumir o valor 2 no tempo 7 é

$$P[X_7 = 2] = p(2; 7)$$

A probabilidade do processo assumir os valores 5 e 2, nos tempos 1 e 7, respectivamente, pode ser escrita como

$$P[X_1 = 5, X_7 = 2] = p(5, 2; 1, 7).$$

Distribuição de ordem k

Exemplo 13

A função de probabilidade de ordem 1 no tempo 2, do passeio aleatório no exemplo 1, é dada por

$$p(x;2) \equiv P[X_2 = x] = \begin{cases} 2p(1-p) & \text{se } x = 0; \\ p^2 & \text{se } x = 2; \\ (1-p)^2 & \text{se } x = -2. \\ 0 & \text{outros casos,} \end{cases}$$

onde $p := P[\{\text{Coroa}\}]$. Primeiramente, note que, com dois arremesso da moeda, a partícula pode dar dois passos para a direita (coroa,coroa), dois para a esquerda (cara,cara), um para a direita e um para a esquerda (coroa,cara), um para a esquerda e um para a direita (cara,coroa).

Distribuição de ordem k

Continuação do exemplo 13

No primeiro caso, a posição final é 2; no segundo é -2; no terceiro e quarto é 0. Portanto, $S_{X_2} = \{-2, 0, 2\}$. Agora,

$$p(2; 2) = P[X_2 = 2] = P[\{\text{Coroa}, \text{Coroa}\}].$$

Supondo que os arremessos são independentes,

$$P[\{\text{Coroa}, \text{Coroa}\}] = P[\{\text{Coroa}\}]P[\{\text{Coroa}\}] = p^2.$$

Logo,

$$p(2; 2) = p^2.$$

Analogamente,

$$p(-2; 2) = (1 - p)^2 \quad \text{e} \quad p(0; 2) = 2p(1 - p).$$

Distribuição de ordem k

Exemplo 14

Vamos obter a função de probabilidade de ordem 1 no tempo 2 da cadeia de Ehrenfest, descrita no exemplo 4, supondo $d = 10$ e que, inicialmente, há 4 bolas na caixa 1. Como $X_0 = 4$, há dois valores possíveis para X_1: 3, se a primeira bola selecionada pertence à caixa 1, e 5, se a primeira bola selecionada pertence à caixa 2. Logo, $S_{X_1} = \{3, 5\}$. A probabilidade de a primeira bola ser da caixa 1 é $4/10$, pois, inicialmente, há 4 bolas na caixa 1. A probabilidade de a primeira bola ser da caixa 2 é $6/10$, pois, inicialmente, há 6 bolas na caixa 2. Portanto,

$$p(x; 1) \equiv P[X_1 = x] = \begin{cases} 4/10, & \text{se } x = 3; \\ 6/10, & \text{se } x = 5. \end{cases}$$

Distribuição de ordem k

Continuação do exemplo 14

Essa é a função de probabilidade de ordem 1 no tempo 1. Há quatro resultados possíveis para X_2:

- $X_2 = 2$, se $X_1 = 3$ e a segunda bola selecionada é da caixa 1;
- $X_2 = 4$, se $X_1 = 3$ e a segunda bola selecionada é da caixa 2;
- $X_2 = 4$, se $X_1 = 5$ e a segunda bola selecionada é da caixa 1;
- $X_2 = 6$, se $X_1 = 5$ e a segunda bola selecionada é da caixa 2 .

Logo, $S_{X_2} = \{2, 4, 6\}$. Pela lei da probabilidade total [2, p. 6],

$$p(x_2; 2) = P[X_2 = x_2] = \sum_{x_1 \in S_{X_1}} P[X_1 = x_1, X_2 = x_2] =$$

$$= P[X_1 = 3, X_2 = x_2] + P[X_1 = 5, X_2 = x_2].$$

Distribuição de ordem k

Continuação do exemplo 14

Pela regra do produto [2, p. 5],

$$P[X_1 = 3, X_2 = x_2] = P[X_1 = 3]P[X_2 = x_2|X_1 = 3] = \\ = \frac{4}{10}P[X_2 = x_2|X_1 = 3].$$

Agora,

$$P[X_2 = x_2|X_1 = 3] = \begin{cases} 3/10 & \text{se } x_2 = 2; \\ 7/10 & \text{se } x_2 = 4. \end{cases}$$

Logo,

$$P[X_1 = 3, X_2 = x_2] = \begin{cases} 4/10 \cdot 3/10 = 12/100, & \text{se } x_2 = 2; \\ 4/10 \cdot 7/10 = 28/100, & \text{se } x_2 = 4. \end{cases}$$

Distribuição de ordem k

Continuação do exemplo 14

Analogamente,

$$\begin{aligned} P[X_1 = 5, X_2 = x_2] &= P[X_1 = 5]P[X_2 = x_2 | X_1 = 5] = \\ &= \frac{6}{10} P[X_2 = x_2 | X_1 = 5] \end{aligned}$$

e

$$P[X_2 = x_2 | X_1 = 5] = \begin{cases} 5/10 & \text{se } x_2 = 4; \\ 5/10 & \text{se } x_2 = 6. \end{cases}$$

Logo,

$$P[X_1 = 5, X_2 = x_2] = \begin{cases} 6/10 \cdot 5/10 = 30/100, & \text{se } x_2 = 4; \\ 6/10 \cdot 5/10 = 30/100, & \text{se } x_2 = 6. \end{cases}$$

Distribuição de ordem k

Continuação do exemplo 14

Temos que

$$p(x_1, x_2; 1, 2) = \begin{cases} 12/100, & \text{se } x_1 = 3 \text{ e } x_2 = 2; \\ 28/100, & \text{se } x_1 = 3 \text{ e } x_2 = 4; \\ 30/100, & \text{se } x_1 = 5 \text{ e } x_2 = 4; \\ 30/100, & \text{se } x_1 = 5 \text{ e } x_2 = 6. \end{cases}$$

Essa é a função de probabilidade de ordem 2 nos tempos 1 e 2 da cadeia de Ehrenfest sendo considerada aqui. Finalmente, a função de probabilidade de ordem 1 no tempo 2 é dada por

$$p(x_2; 2) = P[X_2 = x_2] = \begin{cases} 12/100, & \text{se } x_2 = 2; \\ 58/100, & \text{se } x_2 = 4; \\ 30/100, & \text{se } x_2 = 6. \end{cases}$$

Distribuição de ordem k

Exemplo 15

Seja $\{X(t), t \geq 0\}$ o processo estocástico definido como

$$X(t) = e^Y t,$$

onde $Y \sim U(0,1)$. Vamos calcular a função de densidade de ordem 1 no tempo 1 desse processo. Primeiro, calcularemos a função de distribuição. Em seguida, calcularemos a função de densidade. Note que

$$F(x;t) = P[X(t) \leq x] = P[e^Y t \leq x] = P[Y \leq \ln(x/t)]$$

Como $Y \sim U(0,1)$, temos que

$$P[Y \leq y] = \begin{cases} y & \text{se } 0 < y < 1; \\ 0 & \text{caso contrário.} \end{cases}$$

Distribuição de ordem k

Continuação do exemplo 15

Logo,

$$F(x;t) = \begin{cases} \ln(x/t) & \text{se } 0 < \ln(x/t) < 1; \\ 0 & \text{caso contrário.} \end{cases}$$

Agora, $\ln(x/t) > 0 \Rightarrow x > t$ e $\ln(x/t) < 1 \Rightarrow x < te$. Portanto,

$$F(x;t) = \begin{cases} \ln(x/t) & \text{se } t < x < te; \\ 0 & \text{caso contrário.} \end{cases}$$

Sabemos que $f(x;t) = \frac{\partial}{\partial x} F(x;t)$. Segue-se que

$$f(x;t) = \begin{cases} 1/x & \text{se } t < x < te; \\ 0 & \text{caso contrário.} \end{cases}$$

Referências I

[1] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to stochastic processes*, Houghton Mifflin, Boston, MA, EUA, 1972.

[2] M. Lefebvre, *Applied stochastic processes*, Springer, New York, NY, EUA, 2007.